Technical Note No. 1
Control Theory Group

THE UNIVERSITY OF TENNESSEE

DEPARTMENT OF ELECTRICAL ENGINEERING

METHOD OF ADJOINT SYSTEMS AND ITS ENGINEERING APPLICATIONS

by

J. C. Hung

Prepared for National Aeronautics and Space Administration under Contract NAS8-11183

October 10, 1964

Technical Note No. 1

Control Theory Group

METHOD OF ADJOINT SYSTEMS AND ITS ENGINEERING APPLICATIONS

by

J. C. Hung

Prepared for National Aeronautics and Space

under Contract NAS8-11183

October 10, 1964

THE UNIVERSITY OF TENNESSEE

ELECTRICAL ENGINEERING DEPARTMENT

KNOXVILLE, TENNESSEE

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1. Purpose

Mathematicians have long been paying attention to the adjoint differential equations in their studies of the existence of solutions of differential equations. In recent years, increasing use has been made of the method of adjoint systems to various types of engineering problems, especially in trajectory optimization, two-point boundary value problems, and computer simulation techniques. Although the method of adjoint systems is not very involved in theory, an introduction to this method is difficult to find in extant engineering literature.

This material is prepared to serve as a tutorial account of the method of adjoint systems and its application to engineering problems.

Chapter II provides the fundamentals for the solution of vector differential equations with time-varying coefficients. Chapter III presents the adjoint system in both vector equation form and differential operator form. Several important properties of the adjoint system will also be discussed here. The rest of the chapters deal with the applications of the adjoint method to various engineering problems. A list of references is included at the end of this note.

2. Mathematical Representation of Time-Varying Linear Systems

A linear system is one that obeys the superposition theorem.

Its characteristics are completely specified by the impulse response function $h(t,\tau)$ which gives the system output at time t due to a unit impulse input at time τ. Notice that $h(t,\tau)$ is a function of two variables, t and τ.

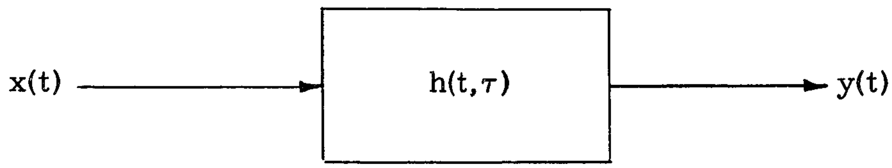

Figure 1

If x(t) is the input of the system and y(t) is the output, we have

$$y(t) = \int_{-\infty}^{\infty} h(t,\tau)\, x(\tau)\, d\tau \tag{1}$$

which amounts to summing up all the responses due to input impulses.

For a physically realizable system

$$h(t,\tau) = 0 \qquad \text{for } \tau > t \tag{2}$$

then (1) becomes

$$y(t) = \int_{-\infty}^{t} h(t,\tau)\, x(\tau)\, d\tau. \tag{3}$$

The concept of impulse response $h(t,\tau)$ as a function of both t and τ is shown in Figure 2. In this figure, τ-axis represents the time when the unit impulse is applied to the system and t-axis represents the time when the output is observed.

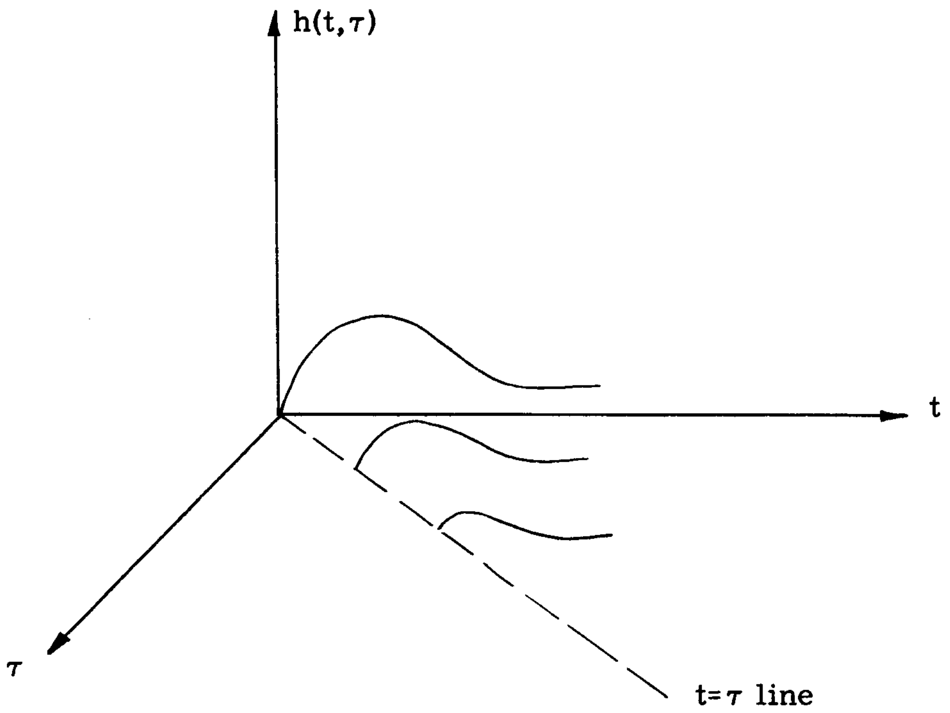

Figure 2

If the system is time invariant,

$$h(t,\tau) = h(t-\tau) = h(t') \tag{4}$$

where $t' = t - \tau$, i.e., the impulse response is a function of a single variable t'. Eq. (1) then reduces to

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)\, x(\tau)\, d\tau \tag{5}$$

and (3) to

$$y(t) = \int_{-\infty}^{t} h(t-\tau)\, x(\tau)\, d\tau \tag{6}$$

which are the usual convolution integrals. By a change of variable, (5) can also be expressed as

$$y(t) = \int_{-\infty}^{\infty} h(\tau)\, x(t-\tau)\, d\tau. \tag{7}$$

The corresponding expression for time-varying system is

$$y(t) = \int_{-\infty}^{\infty} h(t,t-\tau)\, x(t-\tau)\, d\tau. \tag{8}$$

CHAPTER II

SOLUTION OF VECTOR DIFFERENTIAL EQUATIONS WITH TIME-VARYING COEFFICIENTS

In this section we shall solve and discuss the linear differential equation

$$\dot{\underline{x}}(t) = A(t)\,\underline{x}(t) + \underline{b}(t)\,u(t) \qquad \underline{x}(t_o) = \underline{x}_o \tag{1}$$

where $\underline{x}$ and $\underline{b}$ are n-vectors, A is an n x n matrix, and u(t) is a scalar function.

First we shall consider the homogeneous differential equation

$$\dot{\underline{x}}(t) = A(t)\,\underline{x}(t)\ . \tag{2}$$

Property 1: For every given $\underline{x}(t_o) = \underline{x}_o$ there exists a unique solution for the homogeneous differential equation (2).

Property 2: If $\underline{x} = \underline{\phi}(t)$ is a solution of (2), and if $\underline{\phi}(t_o) = 0$ for a certain t_o, then this solution is identically zero, i.e., $\underline{\phi}(t) \equiv 0$.

Property 3: If the vector functions $\underline{\phi}_1(t), \underline{\phi}_2(t), \cdots \underline{\phi}_s(t)$ are the solutions of (2), then any vector function

$$\underline{\phi}(t) = \sum_{i=1}^{s} c_i\,\underline{\phi}_i(t), \tag{3}$$

where c_i are constants, is also a solution of (2).

Property 4: If the system of vector solutions

$$\underline{\phi}_1(t), \underline{\phi}_2(t), \cdots \underline{\phi}_n(t) \tag{4}$$

of (2) are linearly independent for all t, it is called the fundamental system of solutions. A fundamental system of solutions always exists for (2).

Property 5: If (4) is a fundamental system of solutions of (2), then every solution $\underline{\phi}(t)$ of (2) can be expressed as

$$\underline{\phi}(t) = \sum_{i=1}^{n} c_i \underline{\phi}_i(t) \tag{5}$$

where c_i are suitably chosen constants. Except for a few special cases, there is no general method available to obtain $\underline{\phi}_i(t)$ in explicit form.

Property 6: Each vector solution $\underline{\phi}_i$ of (2) can be written in terms of its components as

$$\underline{\phi}_i(t) = \begin{bmatrix} \phi_{1i} \\ \phi_{2i} \\ \vdots \\ \phi_{ni} \end{bmatrix} . \tag{6}$$

The matrix $\underline{\Phi}(t)$ whose column vectors are the vector solutions of the fundamental system of solutions (4), i.e.,

$$\underline{\Phi}(t) = \begin{bmatrix} \phi_{11}(t) & \phi_{12}(t) \cdots\cdots & \phi_{1n}(t) \\ \phi_{21}(t) & \phi_{22}(t) \cdots\cdots & \phi_{2n}(t) \\ \vdots & \vdots & \vdots \\ \phi_{n1}(t) & \phi_{n2}(t) \cdots\cdots & \phi_{nn}(t) \end{bmatrix} , \tag{7}$$

is called the <u>fundamental matrix</u> of (2). The fundamental matrix is nonsingular due to the linear independence of $\underline{\phi}_i(t)$. Note that the fundamental matrix (7) of (2) is not unique.

<u>Property 7:</u> The fundamental matrix $\Phi(t)$ satisfies the matrix differential equation of (2), i.e.

$$\dot{\Phi}(t) = A(t)\,\Phi(t). \tag{8}$$

<u>Property 8:</u> If both $\Phi(t)$ and $\Psi(t)$ are fundamental matrices of (2), then they are related by

$$\Psi(t) = \Phi(t)\,P \tag{9}$$

where P is a constant matrix.

Among the fundamental matrices of (2), the one, denoted by $\Phi(t,t_o)$, which satisfies the initial condition

$$\Phi(t_o,t_o) = I, \tag{10}$$

the unit matrix, is called the <u>transition matrix</u> of (2). Then the solution to the differential equation

$$\underline{\dot{x}}(t) = A(t)\,\underline{x}(t) \tag{2}$$

with the initial condition $\underline{x}(t_o) = \underline{x}_o$ is given by

$$\underline{x}(t) = \Phi(t,t_o)\,\underline{x}_o. \tag{11}$$

<u>Property 9:</u> The transition matrix has the properties

$$\Phi(t_1,t_2)\,\Phi(t_2,t_3) = \Phi(t_1,t_3)\,, \tag{12}$$

$$\Phi(t_1,t_2)^{-1} = \Phi(t_2,t_1). \tag{13}$$

Now let us find the solution of the non-homogeneous differential equation

$$\underline{\dot{x}}(t) = A(t)\,\underline{x}(t) + \underline{b}(t)\,u(t) \qquad \underline{x}(t_o) = \underline{x}_o \tag{14}$$

which is the same as (1). The complementary solution of (14) is the solution of the homogeneous part of (14), which is

$$\underline{x}(t) = \Phi(t,t_o)\,\underline{k} \tag{15}$$

where $\underline{k}$ is a constant vector to be determined. To obtain the particular solution of (14), using the method of variation of parameters, let the particular solution be

$$\underline{x}(t) = \Phi(t,t_o)\,\underline{c}(t), \tag{16}$$

then

$$\underline{\dot{x}} = \dot{\Phi}\,\underline{c} + \Phi\,\underline{\dot{c}} = A\Phi\,\underline{c} + \Phi\,\underline{\dot{c}} = A\underline{x} + \Phi\,\underline{\dot{c}}\ . \tag{17}$$

Comparing (17) to (14),

$$\Phi\,\underline{\dot{c}} = \underline{b}\,u$$

therefore

$$\underline{c}(t) = \int_{t_o}^{t} \Phi^{-1}(\tau,t_o)\,\underline{b}(\tau)\,u(\tau)\,d\tau. \tag{18}$$

The form of the complete solution is therefore

$$\underline{x}(t) = \Phi(t,t_o)\,\underline{k} + \Phi(t,t_o)\int_{t_o}^{t} \Phi^{-1}(\tau,t_o)\,\underline{b}(\tau)\,u(\tau)\,d\tau. \tag{19}$$

Using the initial condition $\underline{x}(t_o) = \underline{x}_o$, we find

$$\underline{k} = \underline{x}_o \ . \tag{19}$$

Hence the complete solution is

$$\underline{x}(t) = \Phi(t,t_o)\underline{x}_o + \Phi(t,t_o) \int_{t_o}^{t} \Phi^{-1}(\tau,t_o)\, \underline{b}(\tau)\, u(\tau)\, d\tau. \tag{20}$$

If the initial condition $\underline{x}_o = 0$,(20) becomes the <u>zero-state-response</u> of (14), which is, with the aid of (12),

$$\underline{x}(t) = \int_{t_o}^{t} \Phi(t,\tau)\, \underline{b}(\tau)\, u(\tau)\, d\tau. \tag{21}$$

When initial condition $\underline{x}_o = 0$ and when $u(\tau) = \delta(\tau - \alpha)$, (20) gives the <u>impulse response</u> of the system (14) in response to a unit impulse applied at time α, which will be denoted by $h(t,\alpha)$. From (20)

$$\underline{h}(t,\alpha) = \Phi(t,\alpha)\, \underline{b}(\alpha). \tag{22}$$

CHAPTER III

THE ADJOINT SYSTEMS

1. The Adjoint Systems in Vector Equation Form

Two systems represented by the linear differential equations

$$\dot{\underline{x}} = A(t)\,\underline{x} \qquad\qquad A(t) \text{ --- real} \tag{1}$$

$$\dot{\underline{y}} = -A^*(t)\,\underline{y}\,, \tag{2}$$

where A^* is the transposed matrix of A, are said to be adjoint to one another.

A^* is often called the adjoint operator of the operator A. The operator and the adjoint operator have the property

$$\langle A\underline{x}, \underline{y}\rangle = \langle \underline{x}, A^*\underline{y}\rangle\ . \tag{3}$$

Theorem: Let $\underline{x}(t)$ and $\underline{y}(t)$ be any solutions of (1) and (2), respectively, then the inner product of the two solutions

$$\langle \underline{x}(t), \underline{y}(t)\rangle \equiv \text{constant}. \tag{4}$$

If $\Phi(t,\tau)$ and $\Psi(t,\tau)$ are the transition matrices of (1) and (2), respectively, then

$$\Psi^*(t,\tau)\,\Phi(t,\tau) \equiv I. \tag{5}$$

The converse is also time.

Proof: $\Phi(t,\tau)$ and $\Psi(t,\tau)$ have the properties (from (8) and (10) of Chapter II) that

$$\frac{d}{dt}\,\Phi(t,\tau) = A(t)\,\Phi(t,\tau), \qquad \Phi(\tau,\tau) = I$$

$$\frac{d}{dt}\,\Psi(t,\tau) = -A^*(t)\Psi(t,\tau), \qquad \Psi(\tau,\tau) = I.$$

Using these properties, we have

$$\frac{d}{dt}\left[\Psi^*(t,\tau)\,\Phi(t,\tau)\right] = \frac{d}{dt}\left[\Psi^*(t,\tau)\right]\Phi(t,\tau) + \Psi^*(t,\tau)\left[\frac{d}{dt}\,\Phi(t,\tau)\right]$$

$$= -\Psi^*(t,\tau)\,A(t)\,\Phi(t,\tau) + \Psi^*(t,\tau)\,A(t)\,\Phi(t,\tau) = 0$$

therefore

$$\Psi^*(t,\tau)\,\Phi(t,\tau) = \text{constant.}$$

But

$$\Psi^{-1}(\tau,\tau)\,\Phi(\tau,\tau) = I.$$

So

$$\Psi^*(t,\tau)\,O(t,\tau) \equiv I$$

which is (5).

From Chapter II, we see that the solution of (1) and (2) are

$$\left.\begin{aligned} \underline{x}(t) &= \Phi(t,\tau)\,\underline{x}(\tau) \\ \underline{y}(t) &= \Psi(t,\tau)\,\underline{y}(\tau) \end{aligned}\right\} \qquad (6)$$

where $\underline{x}(\tau)$ and $\underline{y}(\tau)$ are constants. Therefore, by using (5)

$$\langle \underline{x}(t), y(t) \rangle = \underline{x}^*(t)\,\underline{y}(t)$$

$$= \underline{x}^*(\tau)\,\Phi^*(t,\tau)\,\Psi(t,\tau)\,\underline{y}(\tau)$$

$$= \underline{x}^*(\tau)\,\underline{y}(\tau) = \text{constant}$$

which is (4). Q.E.D.

Remark: (5) implies that the set $\{\underline{\psi}_i\}$ of the column vectors of Ψ forms a reciprocal basis of the basis $\{\underline{\phi}_j\}$ which is the set of column vectors of Φ.

The importance of the adjoint systems is due to the following facts:

1. The nature of the solution of the homogeneous adjoint equations can give the information about the existence of a solution of the non-homogeneous equations from which the adjoint equations came.
2. The important relation between a system and its adjoint system is indispensable for the study of boundary value problems.
3. For a time-varying system, a property of the adjoint system enables us to plot the impulse response of the original system in one computer run.

2. An Important Relationship

Consider a system represented by the differential equation

$$\underline{\dot{x}}(t) = A(t)\,\underline{x}(t) + \underline{b}(t)\,f(t) \tag{7}$$

whose zero-state-response and vector impulse response are, respectively,

$$\underline{x}(t) = \int_{t_o}^{t} \Phi(t,\tau)\, \underline{b}(\tau)\, d\tau \tag{8}$$

$$h(t,\tau) = \Phi(t,\tau)\, \underline{b}(\tau). \tag{9}$$

If the r^{th} variable of the vector $\underline{x}(t)$ is the system output, the impulse response of the system is then

$$h_r(t,\tau) = \sum_j \phi_{rj}(t,\tau)\, b_j(\tau). \tag{10}$$

Next, let us form the adjoint system represented by the adjoint differential equation

$$\dot{\underline{z}}(t) = -A^*(t)\, \underline{z}(t) + \underline{c}(t)\, g(t) \tag{11}$$

whose zero-state-response of (11) is

$$\underline{z}(t) = \int_{t_o}^{t} \Psi(t,\tau)\, \underline{c}(\tau)\, g(\tau)\, d\tau$$

$$= \int_{t_o}^{t} \Phi^*(\tau,t)\, \underline{c}(\tau)\, g(\tau)\, d\tau.$$

Taking the transpose of the last equation,

$$\underline{z}^*(t) = \int_{t_o}^{t} g(\tau)\, \underline{c}^*(\tau)\, \Phi(\tau,t)\, d\tau. \tag{12}$$

Now, for a particular $h_r(T,\tau)$ of (10) having fixed r and T, let

$$g(\tau) = \delta(\tau - T) \tag{13}$$

$$\underline{c}^*(\tau) = \left[\delta_{r1}, \ \delta_{r2}, \ \cdots, -\delta_{rr}, \ \cdots, \ \delta_{rn}\right] \tag{14}$$

and define

$$(\text{Output of the adjoint system}) = \left\langle \underline{z}(t), \underline{b}(t) \right\rangle = \underline{z}^*(t)\, \underline{b}(t). \tag{15}$$

Then the impulse response of its adjoint system at time t due to an input at time T is, denoted by $k_r(t,T)$,

$$\begin{aligned} k_r(t,T) &= \underline{z}^*(t)\, \underline{b}(t) \\ &= -\underline{c}^*(T)\, \Phi(T,t)\, \underline{b}(t) \\ &= -\sum_j \phi_{rj}(T,t)\, b_j(t). \end{aligned} \tag{16}$$

Notice that the form of $\underline{c}^*(\tau)$ as shown in (14) indicates that a negative unit-impulse is applied to the r^{th} differential equation of (11) (to the input of the r^{th} integrator).

Comparing (16) and (10), we have the following important relationship

$$h_r(t,\tau) = -k_r(\tau, t). \tag{17}$$

In summary, (17) implies that the impulse response of a system at time t due to an input at time τ can be obtained from the impulse response of its adjoint system at time τ due to an input at time t, provided that the output of the adjoint system is defined by (15) and the input to the adjoint system is chosen according to (14).

3. Adjoint System in Differential Operator Form

Consider the linear differential equation

$$Ly = p^3 y + a_1(t)p^2 y + a_2(t)py + a_3(t)y = u(t) \tag{18}$$

where $p = \frac{d}{dt}$. Its differential operator is

$$L = p^3 + a_1(t)p^2 + a_2(t)p + a_3(t). \tag{19}$$

Expressing (18) in normal vector differential equation form,

$$\dot{\underline{y}} = A(t)\, \underline{y} + \underline{b}(t)\, u(t) \tag{20}$$

where

$$y = \begin{bmatrix} y_1 = y \\ y_2 = \dot{y} \\ y_3 = \ddot{y} \end{bmatrix}$$

$$A(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3(t) & -a_2(t) & -a_1(t) \end{bmatrix} \tag{21}$$

$$\underline{b}(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The adjoint system of (20) is

$$\dot{\underline{z}} = A^*(t)\, \underline{z} \tag{22}$$

where

$$\underline{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}, \qquad -A^*(t) = \begin{bmatrix} 0 & 0 & a_3(t) \\ -1 & 0 & a_2(t) \\ 0 & -1 & a_1(t) \end{bmatrix} \tag{23}$$

Eliminating z_1 and z_2 from (22) and letting $z_3 = z$, we have the differential equation

$$-p^3z + p^2(a_1(t)z) - p(a_2(t)z) + a_3(t)z = 0 \tag{24}$$

which has a differential operator, denoted by L^*;

$$L^* = -p^3 + p^2a_1(t) - pa_2(t) + a_3(t) \tag{25}$$

where L^* is called the adjoint differential operator.

Remark: The solution of $Ly = 0$ is the first component y_1 of the vector solution of (20); whereas the solution of $L^*z = 0$ is the last component z_3 of (22).

Remark: If the adjoint differential equation (24) has the non-homogeneous term, i.e. if

$$-p^3z + p^2(a_1(t)z) - p(a_2(t)z) + a_3(t)z = v(t) \tag{26}$$

the corresponding vector differential equation becomes

$$\dot{\underline{z}} = -A^*(t)z + \underline{c}(t)\, v(t) \tag{27}$$

where

$$\underline{c}(t) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \tag{28}$$

Comparing $\underline{b}(t)$ of (21) and $\underline{c}(t)$ of (28), we see that the non-homogeneous term appears in the last equation of the set (20) (original system) whereas it appears in the first equation of the set (27) (adjoint system).

The above discussion can be generalized to any n^{th} order linear differential equation. As a result, we have the following.

To every operator

$$L = a_o(t)\, p^n + a_1(t)\, p^{n-1} + \cdots + a_n(t)$$

$$= \sum_{i=0}^{n} a_{n-i}(t)\, p^i(\cdot)\,, \tag{29}$$

where $a_j(t)$ are real, there is associated a second operator

$$L^* = (-1)^n p^n a_o(t) + (-1)^{n-1}\, p^{n-1} a_1(t) + \cdots + a_n(t)$$

$$= \sum_{i=0}^{n} (-1)^i\, p^i\, (a_{n-i}(t)\,\cdot\,) \tag{30}$$

called adjoint of L. Clearly, for L^* to exist, it is necessary that $a_k(t)$ have continuous derivatives through order n-k.

Remark: For time-invariant systems if the original system is

$$Ly = \sum_{i=0}^{n} a_{n-i}\, p^i y = 0 \tag{31}$$

then the adjoint system is

$$L^* z = \sum_{i=0}^{n} (-1)^i p^i\, (a_{n-i} y) = \sum_{i=0}^{n} (-1)^i a_{n-i}\, p^i y = 0. \tag{32}$$

(32) can also be obtained from (20) by replacing t by -t. Thus, for a time invariant system, its adjoint system is a system when running backward in time gives the original system.

Theorem: (Lagrange's Identity) For each pair of functions x(t) and y(t) having derivatives through order n on the interval considered

$$yLx - xL^{*}y = \frac{d}{dt} B(x,y) \tag{33}$$

where

$$B(x,y) = \sum_{i=1}^{n} \Big\{ (p^{i-1}x)\, a_{n-i}y - (p^{i-2}x)\, p(a_{n-i}y) + \cdots$$

$$\cdots + (-1)^{i-1} x\, p^{i-1}(a_{n-i}y) \Big\} \tag{34}$$

and is called bilinear concomitant of Lx.

Proof: Differentiating (34)

$$\frac{d}{dt} B(x,y) = \sum_{i=1}^{n} \Big\{ (p^{i}x)a_{n-i}y + (p^{i-1}x)\, p(a_{n-i}y)$$

$$- (p^{i-1}x)\, p(a_{n-i}y) - (p^{i-2}x)\, p^{2}(a_{n-i}y)$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots +$$

$$(-1)^{i-1}(px)\, p^{i-1}(a_{n-i}y) + (-1)^{i-1} x\, p^{i}(a_{n-i}y) \Big\}$$

$$= \sum_{i=1}^{n} \left\{ (p^i x)\, a_{n-i} y + (-1)^{i-1} x\, p^i\, (a_{n-i} y) \right\}$$

$$= \sum_{i=1}^{n} \left\{ (p^i x)\, a_{n-i} y - (-1)^{i} x\, p^i (a_{n-i} y) \right\} . \tag{35}$$

Using (29) and (30)

$$yLx - xL^* y = \sum_{i=0}^{n} y\, a_{n-i} (p^i x) - \sum_{i=0}^{n} x (-1)^i p^i\, (a_{n-i} y)$$

$$= \sum_{i=1}^{n} \left\{ (p^i x)\, a_{n-i} y - (-1)^i x\, p^i\, (a_{n-i} y) \right\} \tag{36}$$

which is the same as (35). Q.E.D.

Theorem: Let $h(t, \tau)$ be the impulse response for

$$Lx = f \tag{37}$$

and let $k(t, \tau)$ be the impulse response for

$$L^* y = g \tag{38}$$

then,

$$h(t, \tau) = -k(\tau, t). \tag{39}$$

Proof: Let

$x(t) = h(t, \alpha)$

$y(t) = k(t, \beta)$.

Using Lagrange identity,

$$k(t,\beta)Lh(t,\alpha) - h(t,\alpha)L^{*}k(t,\beta) = \frac{d}{dt} B\left[h(t,\alpha), k(t,\beta)\right].$$

By the definition of h and k

$$Lh = L^{*}k = 0$$

therefore

$$B\left[h(t,\alpha), k(t,\beta)\right] = \text{constant for any } t$$

so

$$B\left[h(t,\alpha), k(t,\beta)\right]_{t=\alpha} = B\left[h(t,\alpha), k(t,\beta)\right]_{t=\beta} \tag{40}$$

when $t = \alpha$

$$\begin{aligned} h(t,\alpha) &= \frac{d}{dt} h(t,\alpha) = \cdots = \frac{d^{n-2}}{dt^{n-2}} h(t,\alpha) = 0 \\ \frac{d^{n-1}}{dt^{n-1}} h(t,\alpha) &= \frac{1}{a_o(\alpha)} \end{aligned} \tag{41}$$

When $t = \beta$

$$\begin{aligned} k(t,\beta) &= \frac{d}{dt} k(t,\beta) = \cdots = \frac{d^{n-2}}{dt^{n-2}} k(t,\beta) = 0 \\ \frac{d^{n-1}}{dt^{n-1}} k(t,\beta) &= \frac{(-1)^n}{a_o(\beta)} \end{aligned} \tag{42}$$

Substituting (41) and (42) into (34) to get $B\left[h,k\right]$, then using (40) gives

$$\frac{1}{a_o(\alpha)} a_o(\alpha) k(\alpha,\beta) = (-1)^{n-1} h(\beta,\alpha) \frac{a_o(\beta)}{a_o(\beta)} (-1)^n$$

therefore

$$h(\beta,\alpha) = -k(\alpha,\beta) \qquad \text{Q.E.D.}$$

Remark: (39) agrees with (17) and states that the impulse response of a system can be obtained from the impulse response of its adjoint system. This property enables us to plot the impulse response of a time-varying styste, as a function of the time of application of the unit impulse, in one run on the computer.

Examples:

(i) $$\dot{y} + ay = \delta(t)$$

$$h(t) = e^{-at}$$

$$h(t-\tau) = e^{-a(t-\tau)} .$$

The adjoint system is

$$-\dot{z} + az = \delta(t)$$

$$k(t) = -e^{+at} .$$

Notice that

$$-k(\tau-t) = e^{+a(\tau-t)} = e^{-(t-\tau)} = h(t-\tau) .$$

(ii) $$\ddot{y} + 3\dot{y} + y = \delta(t)$$

$$h(t) = e^{-2t} - e^{-t}$$

$$h(t-\tau) = e^{-2(t-\tau)} - e^{-(t-\tau)} .$$

The adjoint system is

$$\ddot{z} - 3\dot{z} + z = \delta(t)$$

$$k(t) = - e^{+2t} + e^{+t} .$$

Notice that

$$-k(\tau-t) = e^{2(\tau-t)} - e^{(\tau-t)} = e^{-2(t-\tau)} - e^{-(t-\tau)} = h(t-\tau) .$$

CHAPTER IV

APPLICATIONS

1. Application of Adjoint Method to Sensitivity Analysis

A general n^{th} order dynamic system can be represented by a set of n differential equations

$$\dot{x}_i = f_i(x_1, \cdots x_n, u_1, \cdots u_r, t) \tag{1}$$
$$x_i(t_o) = x_i^o, \quad i = 1, 2, \cdots n$$

or, in vector form,

$$\underline{\dot{x}} = \underline{f}\,(\underline{x}, \underline{u}, t), \qquad \underline{x}(t_o) = \underline{x}^o \tag{2}$$

where $\underline{x}$ is an n-vector representing the state of the system and $\underline{u}$ is an r-vector representing the system inputs. Notice that (1) or (2) is, in general, nonlinear and non-autonomous as implied by the form of $\underline{f}$.

Let there be a performance index $J\left[x(t)\right]$ of the system, which depends on the system state $\underline{x}$. Our present problem is to find the variation of $J(\underline{x})$ at time $T > t_o$ due to a change of the control vector $\underline{u}(t)$, $t_o \leq t < T$.

Denote the variations of $\underline{x}$ and $\underline{u}$ by $\delta\underline{x}$ and $\delta\underline{u}$, respectively. From (2)

$$\frac{d}{dt}\,(\hat{\underline{x}} + \delta\underline{x}) = \dot{\hat{\underline{x}}} + \delta\underline{\dot{x}} = \underline{f}\,\left(\hat{\underline{x}} + \delta\underline{x},\ \hat{\underline{u}} + \delta\underline{u},\ t\right)$$

Expanding the last equation in a Taylor series and taking only the linear terms, we have

$$\delta\underline{\dot{x}}(t) = A(t)\,\delta\underline{x}(t) + B(t)\,\delta\underline{u}(t), \tag{3}$$

where

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ | & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad \text{evaluated along } \underline{\hat{x}}(t) \tag{4}$$

$$B = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_r} \\ | & & \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_r} \end{bmatrix} \quad \text{evaluated along } \underline{\hat{u}}(t), \tag{5}$$

which is a time-varying linear differential equation governing the dynamics of state variation due to input purterbation.

The solution of (3) is

$$\delta \underline{x}(t) = \Phi(t,t_o)\, \delta \underline{x}(t_o) + \int_{t_o}^{t} \Phi(t,\tau)\, B(\tau)\, \delta \underline{u}(\tau)\, d\tau. \tag{6}$$

For our present problem $\delta x(t_o) = 0$, then, at $t = T$

$$\delta \underline{x}(T) = \int_{t_o}^{T} \Phi(T,\tau)\, B(\tau)\, \delta \underline{u}(\tau)\, d\tau. \tag{7}$$

The variation of the performance index at time T caused by the variation of the state $\underline{x}$ is

$$J(T) = J\left[\hat{\underline{x}}(T) + \delta\underline{x}(T)\right] - J\left[\hat{\underline{x}}(T)\right] .$$

Again, by series expension technique, the linear portion of the last equation is

$$\delta J(T) = \left\langle \underline{c} , \delta\underline{x}(T)\right\rangle \tag{8}$$

with

$$\underline{c} = \begin{bmatrix} \frac{\partial J}{\partial x_1} \\ \vdots \\ \frac{\partial J}{\partial x_n} \end{bmatrix} \quad \text{evaluated at time T} \tag{9}$$

Substituting (7) into (8)

$$\delta J(T) = \left\langle \underline{c} , \int_{t_o}^{T} \Phi(T,\tau)\, B(\tau)\, \delta\underline{u}(\tau)\, d\tau \right\rangle$$

$$= \int_{t_o}^{T} \left\langle \underline{c}, \Phi(T,\tau)\, B(t)\, \delta\underline{u}(\tau)\right\rangle d\tau. \tag{10}$$

In the integral of (10), $\Phi(T,\tau)$ is considered as a function of its second argument τ with T being fixed. However, the defining equation of the state transition matrix

$$\frac{d}{dt}\,\Phi(t,\tau) = A(t)\,\Phi(t,\tau)$$

specifies the change rate of Φ with respect to the first argument. Therefore, in computing (10), we need to compute a large number of $\Phi(T,\tau)$ by integrating

$$\delta \underline{\dot{x}} = A(t)\, \delta \underline{x}, \qquad \delta \underline{x}(\tau) = I$$

from τ to T for $t_o \leq t < T$ (see Figure 3). This procedure is time consuming and inconvenient.

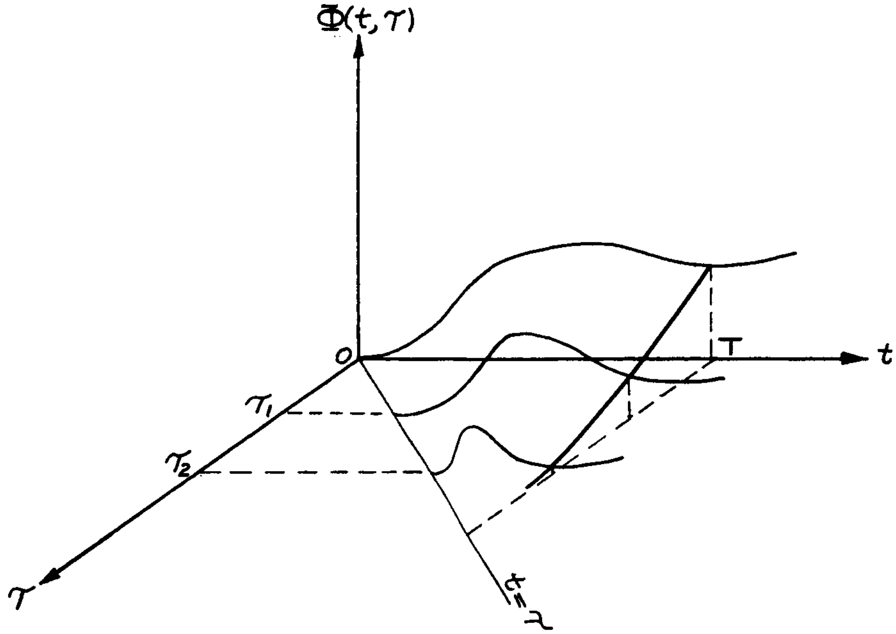

Figure 3

The difficulty can be eliminated by first using the property of adjoint operator, (3) of Chapter III, in (10) to get

$$\delta J(T) = \int_{t_o}^{T} \left\langle B^*(\tau)\, \Phi^*(T,\tau)\, \underline{c},\ \delta \underline{u}(\tau) \right\rangle d\tau$$

and then using property (5) of Chapter III to arrive at

$$J(T) = \int_{t_o}^{T} \left\langle B^*(\tau)\Psi(\tau,T)\,\underline{c},\ \delta\underline{u}(\tau) \right\rangle d\tau \tag{11}$$

where $\Psi(\tau,T)\,\underline{c}$ is the solution of the adjoint system of (3),

$$\dot{\underline{z}}(\tau) = -A^*(t)\,\underline{z}(\tau)\ , \tag{12}$$

with initial condition $\underline{z}(T) = \underline{c}$. Notice that the integration of (12) is going backward in time, since $\tau < T$. Eq. (11) can be integrated with respect to τ in one run.

2. Application of Adjoint Method to Trajectory Optimization

Our next problem is to find the optimum $\underline{u}(t)$ which maximizes the performance index J(T) at time T. Since, in general, the analytic solution is impossible, we have to rely on the high speed computer to work out the answer by an optimum seeking method.

The result obtained in the last section can be used for the optimization. The maximize J(T) is the same as to maximize $\delta J(T)$. Eq. (11) reveals that $\delta J(T)$ is maximized if only if

$$\delta\underline{u}(\tau) = kB^*(\tau)\,\Psi(\tau,T)\,\underline{c} \tag{13}$$

where k is a constant. The constant k is determined by the constraint imposed on $\delta\underline{u}(\tau)$. For example, to preserve accuracy of the computation it is desirable to keep the perturbation $\delta\underline{u}(\tau)$ small enough by requiring

$$\int_{t_o}^{T} \left\langle \delta\underline{u}(\tau),\ \delta\underline{u}(\tau) \right\rangle d\tau \leq r, \text{ a constant.} \tag{14}$$

The procedure required for one iteration is as follows:

(i) Guess $\hat{\underline{u}}(t)$ and find $\hat{\underline{x}}(t)$, $t_o \leq t < T$. Obtain the matrices A(t) and B(t) from $\underline{f}(\underline{x}, u, t)$ by differentiation as shown in (4) and (5), and obtain $\underline{c}$ from $J\left[\underline{x}(T)\right]$ by differentiation as shown in (9).

(ii) Integrate the adjoint differential equation (12) backward to get the quantity $\Psi(t,T)\,\underline{c}$.

(iii) Determine $\delta\underline{u}(t) = kB^*(t)\Psi(t,T)\,\underline{c}$ where k is adjusted to satisfy (14).

(iv) Go back to (i) and use $\hat{\underline{u}} + \delta\underline{u}$ as the new $\hat{\underline{u}}$, then compute the new $\hat{x}(T)$ and new $\delta J(T)$. Keep on the iteration until $\delta J(T)$ is nearly zero.

The above discussion is just one of many possible computation schemes. For example, the dynamics of the system represented by (3), which is a non-homogeneous vector differential equation, can also be represented by a homogeneous vector differential equation (Ref. 5). Furthermore, different iteration procedures can be developed to speed up the solution.

3. Application of Adjoint Method to Two-Point Boundary Value Problem

Consider an n^{th} order dynamic system represented by

$$\dot{x}_i(t) = f_i(x_1 \cdots x_n, t) \qquad i = 1, 2, \cdots n \tag{15}$$

which is, in general, nonlinear and time-varying. The initial values of some of the n state variables and the final values of the rest of the state variables are given, i.e. we have

$$\begin{aligned} &x_1(t_o) = x_1^o, \cdots, x_r(t_o) = x_r^o \\ &x_{r+1}(T) = x_{r+1}^T, \cdots, x_n(T) = x_n^T \end{aligned} \tag{16}$$

(They can be in any order)

Let us call this the split boundary conditions. The problem is to find the solution of (15) satisfying (16). As we know, the analytic solution for (15) is impossible, in general; therefore we are looking for an efficient computer solution.

The usual method of computer solution is to integrate (15) from one end of its trajectory to the other. This requires the knowledge of the values of all the state variables either at the initial end or the final end of the trajectory. In the case of split boundary conditions, the unknown initial values must be guessed and then the integration performed. The guessing and integration need to be done over and over, until the results at the final end meet the given final values. The entire procedure is extremely time consuming and inefficient.

By using the property of adjoint systems, we can convert the two-value boundary problem to an initial value problem, thus allowing a solution to be obtained in one computer run. The method is in the following steps.

<u>Step 1:</u> Guess and/or estimate $x_{r+1}(t_o), \cdots x_n(t_o)$ and denote them by $x^*_{r+1}(t_o), \cdots x^*_n(t_o)$.

<u>Step 2:</u> Integrating (15) from t_o to T, using $x_1(t_o), \cdots,$ $x_o(t_o)$, $x^*_{r+1}(t_o), \cdots x^*_n(t_o)$ as initial conditions, to get $x^*_{r+1}(T), \cdots,$ $x^*_n(T)$.

<u>Step 3:</u> Compute

$$\left.\begin{array}{l} \delta x_{r+1}(T) = x_{r+1}(T) - x^*_{r+1}(T) \\ \quad\vdots \\ \delta x_n(T) = x_n(T) - x^*_n(T) \end{array}\right\} \tag{17}$$

with $x_i(T)$ being considered as the deviation of the state variables at the final end due to a perturbation of initial conditions. The dynamics of $\delta x_i(t)$ along the trajectory is governed, to a first approximation, by

$$\frac{d}{dt}\,\delta x_i(t) = \sum_{j=1}^{n} a_{ij}(t)\,\delta x_j(t), \quad i = 1, \cdots n \tag{18}$$

where $a_{ij}(t) = \dfrac{\partial f_i}{\partial x_j}$ evaluated at $\underline{x}^*$. The adjoint system of (18) is

$$\dot{z}_i = -\sum_{j=1}^{n} a_{ji}(t)\, z_j\,, \qquad i = 1, \cdots n. \tag{19}$$

The vectors $\delta \underline{x}$ and $\underline{z}$, representing δx_i and z_i, satisfy the relation (Chapter III, Equation (4))

$$\frac{d}{dt} \left\langle \delta \underline{x}, \underline{z} \right\rangle = 0 \tag{20}$$

or, more explicitly,

$$\frac{d}{dt} \sum_{i=1}^{n} z_i \, \delta x_i = 0 \ . \tag{20'}$$

Integrating (20) from t_o to T yields

$$\sum_i z_i(T) \delta x_i(T) - \sum_i z_i(t_o) \, \delta x_i(t_o) = 0 \, . \tag{21}$$

In this equation

$$\begin{cases} x_i(t_o) = 0 \quad \text{for} \quad i = 1 \cdots r \\ x_i(t_o) \text{ are unknown for } i = r+1, \cdots, n \\ x_i(T) \text{ are unknown for } i = 1, \cdots, r \\ x_i(T) \text{ are known from (17).} \end{cases}$$

The values of either $z_i(t_o)$ or $z_i(T)$, but not both, are free, and we want to assign them in such a way that results in an efficient computation.

Step 4: Choosing

$$z_i(T) = 0, \quad \text{for} \quad i = 1 \cdots r \tag{22}$$

(21) becomes

$$\sum_{i=r+1}^{n} z_i(T) \ \delta x_i(T) - \sum_{i=1}^{n} z_i(t_o) \ \delta x_i(t_o) = 0. \tag{23}$$

Then, assign

$$z_i(T) = \text{any convenient value}, \quad \text{for } i = r + 1, \cdots n. \tag{24}$$

Integrate the adjoint equation backward from T to t_o, using the boundary values (22) and (24), to achieve $z_i(t_o)$, $i = 1, \cdots, n$.

Step 5: Since there are n-r unknowns, $\delta x_{r+1}(t_o), \cdots, \delta x_n(t_o)$, n-r simultaneous algebraic equations in the form of (23) are needed for the solution. Therefore n-r different set of (24) must be chosen to produce n-r different sets of $z_i(t_o)$, $i = 1, \cdots, n$.

Step 6: Adjust the new estimate to be

$$\left[\overset{*}{x}_i (t_o) \right]_{new} = \left[\overset{*}{x}_i (t_o) \right]_{old} + \delta x_i(t_o) \quad i = r + 1, \cdots, n \tag{25}$$

Step 7: Iterate the procedure until $\delta x_i(T)$ of (17) is small enough to be tolerable.

The above discussion assumed that the final time T is known. In the case where T is not known, its value again needs to be guessed. Let τ be the gesses T, the let

$$\delta T = T - \tau. \tag{26}$$

The deviations from the given end conditions at τ, are not $\delta x_i(T)$, but symbolically $\Delta x_i(\tau)$. Two kinds of deviations are related, to a first approximation, by

$$\delta x_i(T) = \Delta x_i(\tau) - \dot{x}_i(\tau)\, dT \qquad i = 1, \cdots n. \tag{27}$$

Substituting (27) into (23) yields

$$\sum_i z_i(\tau)\left[\Delta x_i(\tau) - \dot{x}_i\, \delta T\right] - \sum_i z_i(t_o)\, \delta x_i(t_o) = 0. \tag{28}$$

Since one more unknown, δT, needs to be found, one more set of $z_i(T)$ must be assigned to get a total of n-r+1 equations in the form of (28). The correction in the value of final time for the next iteration is

$$\left[\tau\right]_{new} = \left[\tau\right]_{odd} + \delta T. \tag{29}$$

4. Application of Adjoint Method to Computer Simulation

Suppose that we want to simulate the impulse response $h(T,\tau)$ of a time-varying linear system as a function of τ, the application time of the input. The defining equation of $h(T,\tau)$, (7) or (18) of Chapter III, specifies that the derivatives of $h(t,\tau)$ be taken with respect to its first argument t. Therefore the direct simulation of $h(T,\tau)$ requires a large number of computer runs, each for a different value of τ.

To make the task simple, we use the property

$$-h(t,\tau) = k(\tau,t) \qquad \tau \leq t \tag{30}$$

where k is the impulse response of the adjoint system. Since τ is the first argument of k, we can obtain $k(\tau,t)$ as a function of τ in one computer run.

It should be noted that (30) states that in simulating $k(\tau,t)$ time τ is running backward starting from t. However, the physical realizability imposes the condition

$$T > \tau \text{ for } h(T,\tau)$$
$$\tau > T \text{ for } k(\tau,T) \tag{31}$$

Furthermore, in practice, it is more natural to make the starting time, when the unit impulse is applied, zero rather than T. These two considerations can be handled by replacing t by T-t in $k(\tau,t)$, since

$$\begin{aligned} h(T,\tau) &= \Big[\, h(t,\tau) \,\Big]_{t=T} \\ &= \Big[h(T-t,\tau)\Big]_{t=0} \\ &= \Big[-k(\tau,T-t)\Big]_{t=0} \end{aligned} \tag{32}$$

where t is time of application of input impulse to the adjoint system.

The adjoint method of computer simulation is often applied to statistical problems. Frequently, we want to use the analog computer to analyze the mean-square error of a given time-varying system having a stationary random input x(t). Let $h(t,\tau)$ represent the impulse response between the system input and error output as shown in Figure 4. The power spectral

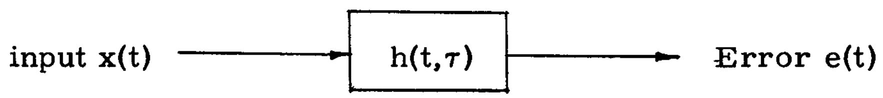

Figure 4

density function of x(t) is $\Phi_{xx}(s)$ which can be decomposed into

$$\Phi_{xx}(s) = G\bar{G} \tag{33}$$

where G = G(s) has only LHP poles and zeros and $\bar{G}$ has only RHP ones. Let the inverse Laplace transform of G be g(t). Then Figure 4 is equivalent to Figure 5, in which the function $h_1(t,\tau)$ represents the tandem combination of g(t) and $h(t,\tau)$.

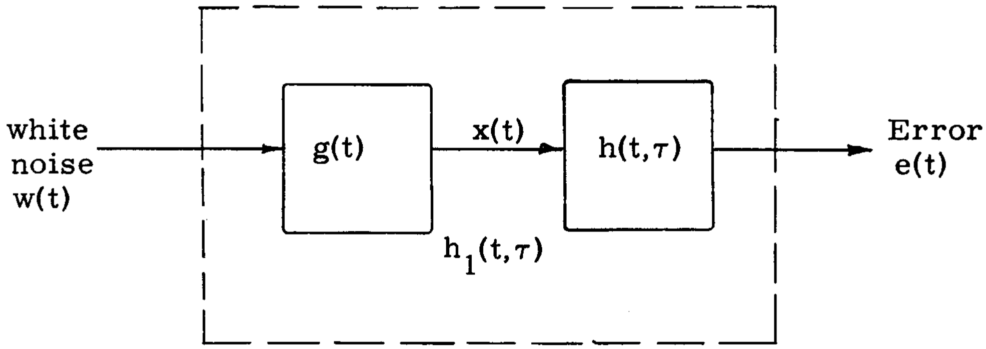

Figure 5

The auto-correlation function for white noise is $\delta(t)$ and the mean-square error of the system is

$$e(t)^2 = \int_{-\infty}^{t}\int_{-\infty}^{t} h_1(t,\tau_1)h_1(t,\tau_2)\ \delta(\tau_1-\tau_2)\ d\tau_1\ d\tau_2$$

$$= \int_{-\infty}^{t} h_1(t,\tau_1)^2\, d\tau_1\ . \tag{34}$$

By letting $\tau_1 = t - \tau$,

$$e(t)^2 = -\int_{0}^{\infty} h_1(t,t-\tau)^2\ d\tau = \int_{0}^{\infty} k_1(t-\tau,t)^2\ d\tau \tag{35}$$

where k_1 is the adjoint of h_1.

A general rule, formulated by Laning and Battin (Ref. 8), for similating the adjoint of a linear time-varying system is given below

> "An analog computer program of the original system whose impulse response we desire is formed using only integrators, summers, and time-varying scale factors. Then the roles of inputs and outputs are reversed for each of the three basic elements, and the argument t in the time-varying coefficients is replaced by T-t. If the output of the r^{th} integrator is the output of the original system, then the impulse input in the modified system should be applied to the input of the r^{th} integrator and the appropriate output recorded".

5. General Remark

We have discussed in this chapter a number of applications of the adjoint method to engineering problems. An efficient use of the method involves not only a knowledge of the properties of adjoin systems, but also the art of computation scheme. Improvements can still be made in the scheme of computation to achieve a still more efficient way of solution.

REFERENCES

1. Bennett, R. R.: Analog Computing Applied to Noise Studies, Proc. IRE, Vol. 41, pp 1509-1513, (1953).
2. Bliss, G. A.: "Mathematics for Exterior Ballistics", John Wiley and Sons, Inc., New York, 1944.
3. Coddington, E. A. and N. Levinson: "Theory of Ordinary Differential Equations", McGraw-Hill Book Co., Inc., New York, 1955.
4. Friedman, B.: "Principles and Techniques of Applied Mathematics", John Wiley and Sons, Inc., New York, 1956.
5. Gavrilovic, M., R. Petrovic and D. Siljak: Adjoint Method in the Sensitivity Analysis of Optimal Systems, J. Franklin Institute, Vol. 276, No. 7, pp 26-38, (1963).
6. Halmos, P. R.: "Finite Dimensional Vector Spaces", 2nd ed., D. Van Nostrand Co., Inc., Princeton, N. J., 1958.
7. Jurovics, S. A. and J. E. McIntyre: The Adjoint Method and Its Application to Trajectory Optimization, ARS Journal, September, (1962).
8. Laning, J. H., Jr., and R. H. Battin: "Random Processes in Automatic Control", McGraw-Hill Book Co., Inc., New York, 1956.
9. Tsien, H. S.: "Engineering Cybernetics", McGraw-Hill Book Co., Inc., New York, 1954.
10. Zadeh, L. A. and C. A. Desoer, : "Linear System Theory", McGraw-Hill Book Co., Inc., New York, 1963.